

Lex4all

Generating Pronunciation
Lexicons for Small-Vocabulary
ASR in LRLs

Our goal

- Pronunciation Lexicon

```
<lexeme>
  <grapheme>ookan</grapheme>
  <phoneme>O L L A0 O</phoneme>
  <phoneme>O L A0 O</phoneme>
  <phoneme>O O H A0 O</phoneme>
</lexeme>
<lexeme>
  <grapheme>meji</grapheme>
  <phoneme>M EI JH I</phoneme>
  <phoneme>M E JH I</phoneme>
  <phoneme>M EI JH JH I</phoneme>
</lexeme>
```

Outline

- Background
- Idea/Components
- Frontend
 - I/O
 - GUI
- Backend
 - ASR Platform
 - Algorithm
- Extensions

Background

- ASR Components
 - acoustic model $P(\text{phoneme}|\text{sound})$
 - trained by labeled data
 - language model $P(\text{word}|\text{word})$
 - grammar for specific tasks
 - lexicon mapping phoneme sequences to words

Background

- **ASR**
 - data-intensive training
 - exists for common languages
- **LRLs**
 - scarce data
 - ASR almost impossible

Idea

- Usage of existing speech recognizers to build lexicons for LRLs
 - Steps
 - Algorithm selects appropriate phonological representation of words by splitting it into single phonemes
 - Building a lexicon that maps written words to possible pronunciations
 - Lexicon is then passed to a recognition application

Idea

- **Components**
 - interface to existing SR
 - algorithm implementation
 - I/O functionality
 - (user interface)

Frontend

- I/O
 - User input
 - String (orthographical representation)
 - one or more .wav file (word-sound)
 - User output
 - xml file (lexicon), mapping orthographical form to pronunciation proposals

Frontend

- GUI (additional)
 - [gui_preview.html](#)
 - more functionality
 - number of pronunciations per word
 - name
 - location of lexicon

Backend

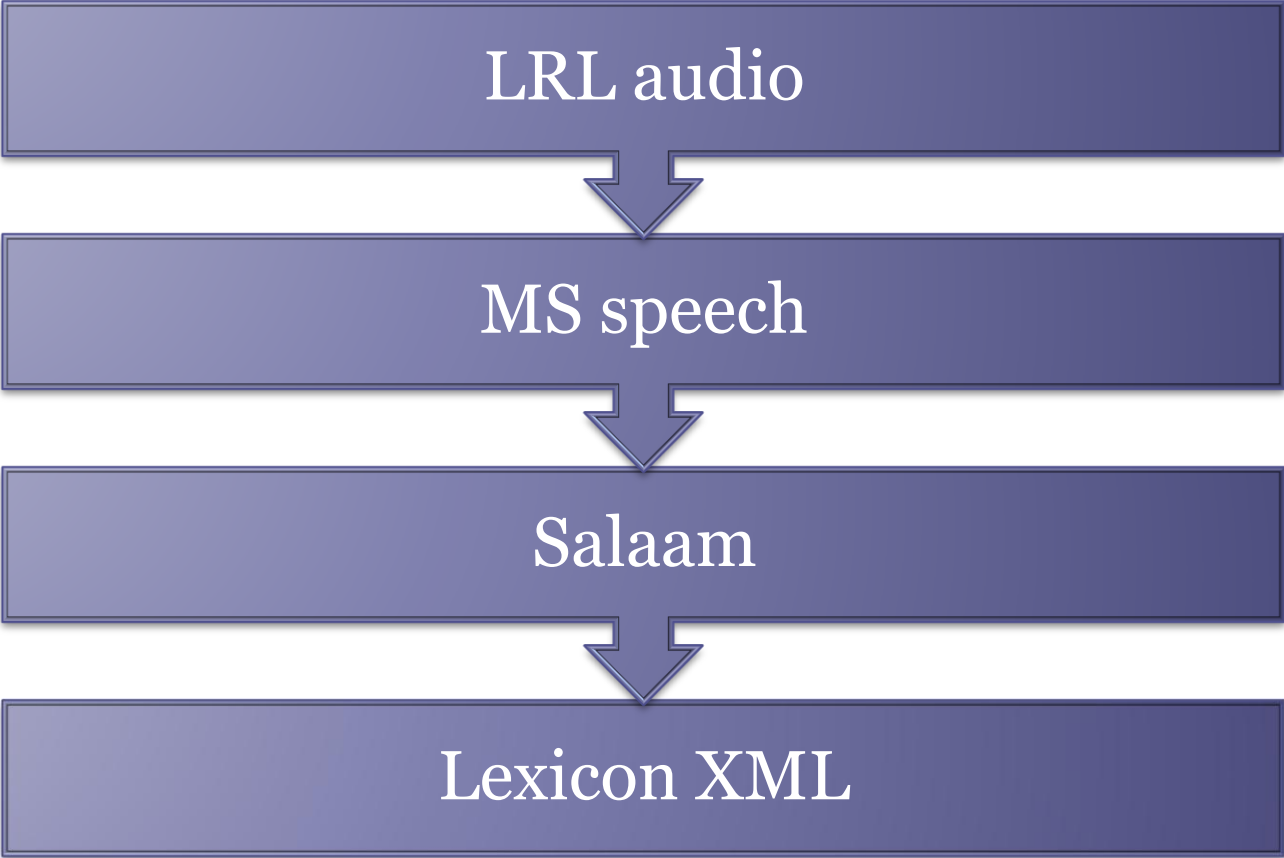
- ASR Platform
 - trade-off between quality and speed
 - MS speech
 - CMU Sphinx
 - Brno phoneme recognizer

Backend

- **Algorithm**
 - basis: CMU Salaam method
 - mimics phoneme recognition using wildcard grammar
 - iterative process, one-by-one recognition
 - disadvantages
 - slow speed -> split grammar, language model

Extensions

- in-app recording
- choice of HRL
 - better recognition results for related languages



References

- Hao Yee Chan and Roni Rosenfeld
 - Discriminative pronunciation learning for speech recognition for resource scarce languages“
- Fang Qiao, Jahanzeb Sherwani, and Roni Rosenfeld
 - Small-vocabulary speech recognition for resource-scarce languages”.